

NVAE: A Deep Hierarchical Variational Autoencoder

Year: 2020 | Citations: 1037 | Authors: Arash Vahdat, Jan Kautz

Abstract

Normalizing flows, autoregressive models, variational autoencoders (VAEs), and deep energy-based models are among competing likelihood-based frameworks for deep generative learning. Among them, VAEs have the advantage of fast and tractable sampling and easy-to-access encoding networks. However, they are currently outperformed by other models such as normalizing flows and autoregressive models. While the majority of the research in VAEs is focused on the statistical challenges, we explore the orthogonal direction of carefully designing neural architectures for hierarchical VAEs. We propose Nouveau VAE (NVAE), a deep hierarchical VAE built for image generation using depth-wise separable convolutions and batch normalization. NVAE is equipped with a residual parameterization of Normal distributions and its training is stabilized by spectral regularization. We show that NVAE achieves state-of-the-art results among non-autoregressive likelihood-based models on the MNIST, CIFAR-10, CelebA 64, and CelebA HQ datasets and it provides a strong baseline on FFHQ. For example, on CIFAR-10, NVAE pushes the state-of-the-art from 2.98 to 2.91 bits per dimension, and it produces high-quality images on CelebA HQ. To the best of our knowledge, NVAE is the first successful VAE applied to natural images as large as 256 \times 256 pixels. The source code is available at this [https URL](https://url).